

НОВАЯ ТЕХНИКА НЕЗАВИСИМОГО ФАКТОРНОГО АНАЛИЗА ВАРИАБЕЛЬНОСТИ СЕРДЕЧНОГО РИТМА

А.В. Мартыненко, А.С. Антонова, А.М. Егоренков

Харьковский национальный университет им. В.Н. Каразина

РЕЗЮМЕ

Данная статья посвящена развитию статистического подхода, основанного на идее разделения регистрируемой композитной на независимые составляющие, формирующиеся соответствующим регуляторным отделом организма. Формальная математическая процедура независимого факторного анализа (ICA) варибельности сердечного ритма (ВСР) позволяет выделить не более трех составляющих в ритмограмме здорового человека, что адекватно отражает физиологические представления об участии систем регуляции в формировании феномена ВСР. Применение ICA анализа ВСР оправданно на временных интервалах от 4 до 15 минут. Выход за указанный временной интервал сопряжен с существенным ухудшением качества выделения формирующих ритмограмму компонент. Оптимальным, с точки зрения качества разделения исходного регистрируемого сигнала ВСР на составляющие, является использование ICA анализа с 5 минутным протоколом регистрации ВСР.

КЛЮЧЕВЫЕ СЛОВА: ритмограмма, варибельность сердечного ритма, независимый факторный анализ

ВВЕДЕНИЕ

Феномен варибельности сердечного ритма (ВСР) был отмечен медиками очень давно: в европейской научной литературе первое упоминание о ВСР связывают с именем Stephen Hales (1733). Однако, действительно широкое развитие и проникновение в медицину методов диагностики, базирующихся на ВСР, приходится на конец прошлого века, что связано с развитием высокоточной цифровой компьютерной электрокардиографии (ЭКГ) и стандартизацией технологии ВСР рабочей группой Европейского Общества Кардиологов и Северо-Американского Общества Стимуляции и Электрофизиологии.

Для исследования варибельности сердечного ритма наиболее часто используются 5-минутный, 15-минутный и 24-часовой протоколы. При анализе ВСР особо ценную информацию дают пространственно-спектральные методы: быстрого преобразования Фурье, ауторегрессионный и др. Важными характеристиками спектра являются его мощность и мощности его отдельных зон (доменов). В 5 и 15 минутных спектрах выделяются 3 характерные зоны: VLF - зона очень низких частот (0.0033 - 0.04 Гц), LF - зона низких частот (0.04 - 0.15 Гц) и HF - зона высоких частот (0.15-0.4 Гц) [1, 6]. При этом физиологически установлена четкая корреляция между мощностью каждого из спектральных доменов и определенным регуляторным отделом организма. На быстрые изменения сердечного ритма (HF) влияет парасимпатическое звено вегетативной нервной регуляции организма; на средние изменения (LF) - симпатическое звено вегетативной нервной регуляции организма, на медленные (VLF) - гуморальное звено нейрогуморальной регуляции организма.

Вместе с тем, стандартный подход не лишен определенных недостатков, связанных, в первую очередь, с формальной техникой выделения в спектре регуляторных доменов. В решении указанной проблемы существенный прогресс может быть достигнут за счет привлечения более сложных математических методов, позволяющих разделить исходную регистрируемую ритмограмму на независимые составляющие. Отметим два направления в подобном неформальном подходе к выделению спектральных-регуляторных доменов в ВСР:

- статистические методы анализа, базирующиеся на общих подходах к анализу временных рядов;
- математическое моделирование, воспроизводящее физиологические особенности регуляторных процессов сердечной деятельности.

Данная статья посвящена развитию статистического подхода, основанного на идее разделения регистрируемой композитной ритмограммы на независимые составляющие, формирующиеся соответствующим регуляторным отделом организма. Очевидно, что применение для этих целей таких известных методов как факторный анализ (РСА), использующих линейную декорреляцию компонент, ограничено и даже неправомерно из-за существенной нелинейности наблюдаемого явления - варибельности сердечного ритма. Появление нелинейного метода статистического анализа - независимого факторного (компонентного) анализа (ICA) [2] - и успешное его развитие для решения сложных медицинских задач (например, анализ электроэнцефалограмм) делает весьма перспективным его применение для анализа ВСР.

МАТЕРИАЛЫ И МЕТОДЫ

Чтобы строго определить применение ИСА для анализа ВСР, используем статистические «скрытые переменные» моделей [3, 4]. Предположим, что мы наблюдаем n линейных смесей x_1, \dots, x_n независимых компонентов (управляющее влияние каждого регуляторного отдела)

$$x_j = a_{j1}s_1 + a_{j2}s_2 + \dots + a_{jn}s_n \text{ для всех } j \quad (1)$$

Здесь опущен временной индекс t . В модели ИСА вместо соответствующего временного отсчета предполагается, что каждая смесь $x_j(t)$, также как каждый независимый компонент s_k - случайная величина. Наблюдаемые значения $x_j(t)$ являются выборкой этой случайной величины (нас не смущает то, что в ВСР в качестве исходной наблюдаемой величины выступает скалярная функция HR, благодаря известной теореме F.Takens [5]). Без потерь общности можно предполагать, что и переменные смеси и независимые компоненты имеют нулевые средние величины. Если это не так, то наблюдаемая переменная x_i может всегда центрироваться, вычитанием из нее выборочного среднего, что приводит к модели с нулевым средним.

Вместо сумм в предыдущем уравнении удобно использовать обозначения с вектор-матрицей. Пусть через \mathbf{x} обозначен случайный вектор, чьи компоненты - смеси x_1, \dots, x_n , и аналогично \mathbf{s} - случайный вектор с компонентами s_1, \dots, s_n . Через \mathbf{A} обозначена матрица с элементами a_{ij} . Обычно малые жирные символы указывают на векторы, а большие жирные символы обозначают матрицы. Все векторы понимаются как векторы-столбцы; таким образом, \mathbf{x}^T (транспонированный \mathbf{x}) является вектор-строкой. Используя эти обозначения с вектор-матрицами, вышеупомянутая модель смешивания запишется как

$$\mathbf{x} = \mathbf{A}\mathbf{s} \quad (2)$$

Иногда мы нуждаемся в столбцах матрицы \mathbf{A} , которые обозначены как \mathbf{a}_j . Модель может также быть записана как

$$\mathbf{x} = \sum_{i=1}^n \mathbf{a}_i s_i \quad (3)$$

Статистическая модель в уравнении 4 называется независимым факторным анализом, или моделью ИСА. Модель - это ИСА порождающая модель, что означает, что она описывает, как наблюдаемые данные сгенерированы процессом смешивания s_i . Независимые компоненты являются скрытыми переменными, что означает, что они не могут непосредственно наблюдаться. Также принято, что матрица смешивания неизвестна. Все, что мы наблюдаем - случайный вектор \mathbf{x} и мы должны с его использованием оценить \mathbf{A} и \mathbf{s} . Это должно быть сделано в насколько

возможно общих предположениях.

Исходная точка для ИСА - самое простое предположение, что компоненты s_i являются статистически *независимыми*. Так же необходимо предположить, что независимый компонент должен иметь *не Гауссово* распределение. Однако в исходной модели мы *не предполагаем* эти распределения известными (если они известны, то проблема значительно упрощается.) Для простоты, также предполагаем, что неизвестная матрица смешивания - квадратная, но это предположение может иногда ослабляться. Тогда, после оценки матрицы \mathbf{A} , можно вычислить ее инверсную матрицу, обозначенную \mathbf{W} , и получить независимый компонент просто как:

$$\mathbf{s} = \mathbf{W}\mathbf{x} \quad (4)$$

ИСА очень близко связано с методом, называемым *слепым разделением источников* (BSS) или слепым разделением сигналов. «Источник» - здесь первоначальный сигнал, то есть независимый компонент. «Слепой» - означает, что мы знаем очень мало, если вообще что ни будь знаем, о матрице смешивания и делаем мало предположений об исходных сигналах. ИСА - один из методов, возможно наиболее широко используемый, для выполнения слепого разделения источников.

Интуитивно говоря, ключ к оценке моделей ИСА - не гауссоидальность. Без негауссоидальности расчет фактически вообще не возможен. В большинстве приложений классической статистической теории принимается, что случайные величины имеют Гауссово распределение, препятствуя, таким образом, любым методам соотносится с ИСА.

Центральная предельная теорема, классический результат теории вероятности, сообщает, что при известных условиях распределение суммы независимых случайных величин стремится к Гауссову распределению. Таким образом, сумма двух независимых случайных величин обычно имеет распределение, которое является ближе к Гауссову чем любая из двух первоначальных случайных величин.

Пусть теперь предполагается, что вектор данных \mathbf{x} распределен согласно модели данных ИСА в уравнении 4, то есть это смесь независимых компонентов. Для простоты, предположим, что все независимые компоненты имеют идентичные распределения. Чтобы оценить один из независимых компонентов, мы рассмотрим линейную комбинацию x_i

$$y = \mathbf{w}^T \mathbf{x} = \sum_i w_i x_i$$

обозначив \mathbf{w} , где \mathbf{w} является вектором, который должен быть опреде-

лен. Если бы \mathbf{w} был одной из строк инверсной матрицы \mathbf{A} , эта линейная комбинация фактически равнялась бы одному из независимых компонентов. Теперь вопрос: Как мы могли бы использовать центральную предельную теорему, чтобы определить \mathbf{w} так, чтобы она равнялась одной из строк инверсии \mathbf{A} ? Практически, мы не можем определять такое \mathbf{w} точно, т.к. мы ничего не знаем относительно матрицы \mathbf{A} , но мы можем найти оценочную функцию, которая даст хорошее приближение.

Чтобы увидеть, как это ведет к основному принципу оценки ICA, сделаем замену переменных, определяя $\mathbf{z} = \mathbf{A}^T \mathbf{w}$. Тогда мы имеем $\mathbf{y} = \mathbf{w}^T \mathbf{x} = \mathbf{w}^T \mathbf{A} \mathbf{s} = \mathbf{z}^T \mathbf{s}$. Таким образом \mathbf{y} - линейная комбинация s_i с весами, данными z_i . Так как сумма двух однородных независимых случайных величин более Гауссова, чем первоначальные переменные, $\mathbf{z}^T \mathbf{s}$ является более Гауссовым, чем любой s_i и становится наименее Гауссовой, когда она фактически равняется одному из s_i . В этом случае, очевидно, только один из элементов z_i матрицы \mathbf{z} является отличным от нуля. (Обратите внимание, что здесь было принято, что s_i имеют идентичные распределения.)

Следовательно, мы можем считать \mathbf{w} вектором, который максимизирует негауссоидальность $\mathbf{w}^T \mathbf{x}$. Такой вектор обязательно соответствовал бы (в преобразованной системе координат) вектору \mathbf{z} , который имеет только один отличный от нуля компонент. Это означает, что $\mathbf{w}^T \mathbf{x} = \mathbf{z}^T \mathbf{s}$ равняется одному из независимых компонентов!

Таким образом, максимизация негауссоидальности $\mathbf{w}^T \mathbf{x}$ дает нам один из независимых компонентов. Фактически, рельеф оптимизации для негауссоидальности в n -мерном пространстве векторов \mathbf{w} имеет $2n$ локальных максимумов, два для каждого независимого компонента, соответствуя s_i и $-s_i$ (подтверждая, что независимые компоненты могут быть оценены с точностью только до мультипликативного знака). Чтобы найти несколько независимых компонентов, мы должны найти все эти локальные максимумы. Это нетрудно, потому что различные независимые компоненты некоррелированы: мы можем всегда ограничивать поиском в пространстве, который дает оценки, некоррелированные с предыдущими. Это соответствует ортогонализации в соответствующем преобразованном пространстве.

Описанный математический метод и алгоритм были реализованы в виде модификации стандартного пакета быстрого ICA для системы MATLAB. В качестве пространственно-спектрального метода анализа ВСР применялось быстрое преобразование Фурье. В исследовании были использованы 24-часовые записи ритмограмм здоровых добровольцев обоего пола. Анализировались участки ритмограммы произвольной длины и произвольного положения внутри мониторинговой записи. Целью исследования было:

- установить адекватность применения математического метода (ICA) анализа variability сердечного ритма физиологическим представлениям о системах регуляции, воплощающих феномен ВСР;
- определить границы применимости ICA анализа ВСР;
- найти оптимальные условия для использования ICA в анализе ВСР.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Пример 8 минутной ритмограммы приведен на рис.1. На рис. 2-4 даны попарно выделенные из исходной ритмограммы с помощью ICA анализа 2, 3 и 4 составляющие (а) и их спектры (б). Спецификой метода является строгое выполнение Пирсоновской декорреляции между разделенными компонентами, что, однако, не обеспечивает декорреляции их спектров. На рис. 5 и 6 даны зависимости от времени (количества ударов сердца) коэффициентов корреляции (*corr*) между спектрами разделенных компонент исходной ритмограммы. Хорошо видно, что применение ICA анализа ВСР ограничено временным диапазоном от 3 до 15 минут, с выделяемым оптимумом в районе 5-7 минут. В табл.1 приведены коэффициенты корреляции для случая выделения в исходной ритмограмме 4-х составляющих. Легко заметить, что введение четвертой компоненты является искусственным: предлагаемая методика определяет высокую степень коррелированности между 3 и 4 сигналами, т.е. невозможность их разделения (рис.4), (табл.1).

Таким образом, ICA анализ ВСР использовался для определения количества независимых переменных, на которых естественным образом раскладывается ритмограмма. Если количество переменных равнялось ожидаемому числу регуляторных систем организма, то есть трем, то каждый из спектров этих сигналов попадал только в одну из областей низко средне или высокочастотную. В то же время если делалась попытка ис-

пользовать четыре или более независимых компонент, то оказывалось, что как минимум две из них попадают в одну и ту же час-

тотную область и более того, их спектры оказывались сильно коррелированы между собой (рис.4).

Таблица 1

Корреляции между спектрами для 4-х составляющих						
Удары сердца	Парные корреляции (<i>corr</i>)					
400	0.027	0.038	0.084	0.156	0.285	0.842
450	0.002	0.046	0.053	0.118	0.278	0.9
500	0.004	0.023	0.1	0.147	0.226	0.869
550	0.061	0.075	0.078	0.14	0.167	0.985
600	0.018	0.044	0.121	0.163	0.174	0.998
650	0.036	0.057	0.15	0.136	0.164	0.994
700	0.04	0.043	0.047	0.254	0.266	0.983
750	0.012	0.02	0.075	0.315	0.324	0.997
800	0.061	0.078	0.234	0.234	0.356	0.996
850	0.016	0.037	0.176	0.186	0.224	0.997
900	0.036	0.063	0.079	0.152	0.254	0.979

Аналогичные результаты получались во всех случаях, когда использовался разложения больше чем на 3 независимые компоненты, - по настоящему независимыми (то есть слабо коррелированными в спектраль-

ной области) оказывались только 3 сигнала. Что подтверждает с одной стороны отсутствие других регуляторных систем организма, а с другой - правомочность предложенного метода анализа.

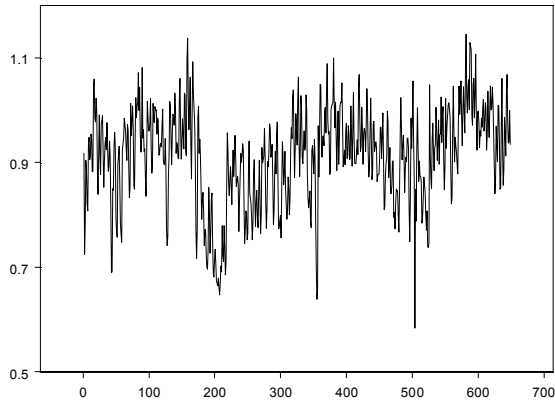


Рис.1а. Исходная ритмограмма (8 минутная запись)

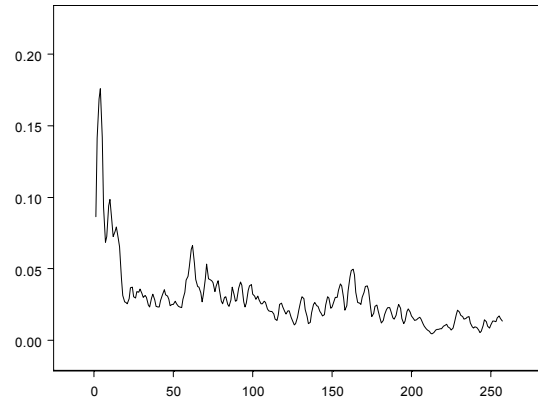


Рис.1б. Спектр исходной ритмограммы

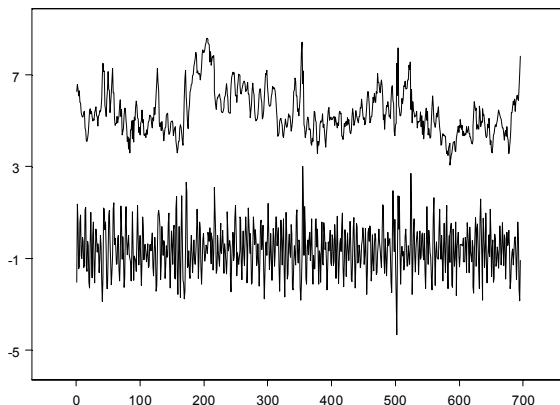
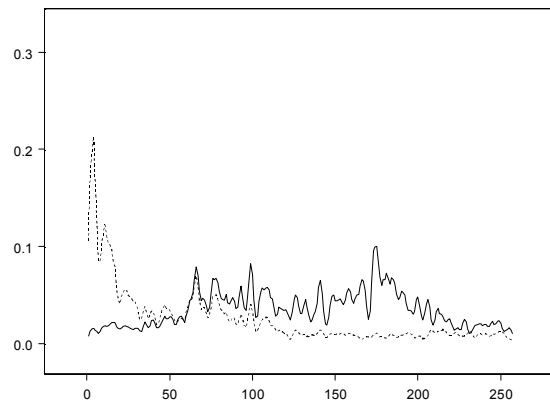


Рис.2а. Ритмограмма, разделенная на 2 составляющие

Рис.2б. Спектр составляющих (коэффициент корреляции между спектрами $corr=0.1$)

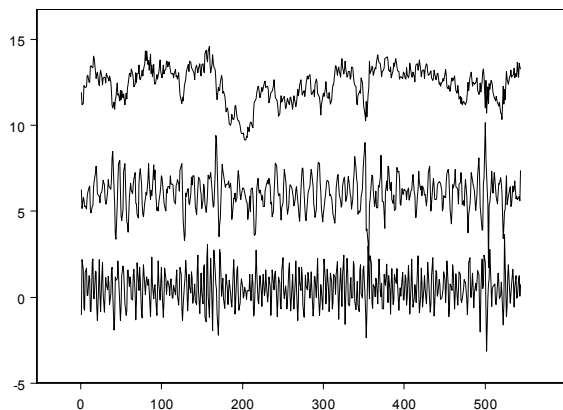


Рис. 3а. Ритмограмма, разделенная на 3 составляющие

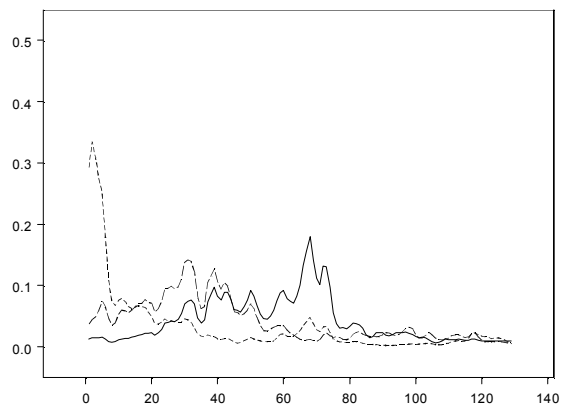


Рис. 3б. Спектр составляющих ($corr=0.34;0.24;-0.04$)

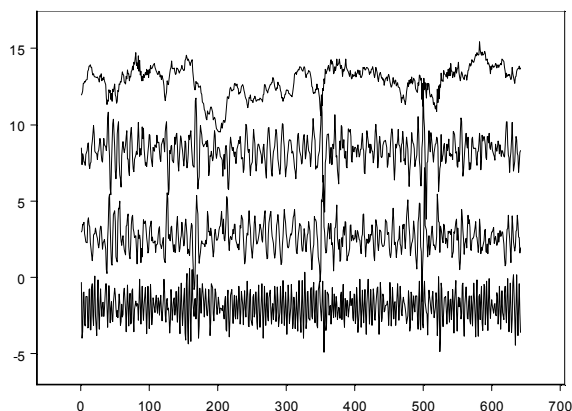


Рис. 4а. Ритмограмма, разделенная на 4 составляющие

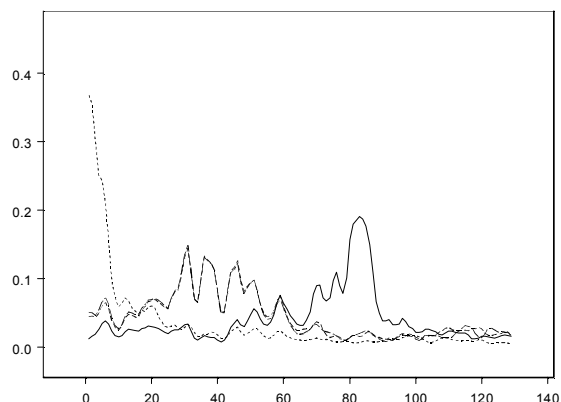


Рис. 4б. Спектр составляющих ($corr=0.99;0.31;0.3;0.28;0.26;0.06$)

Важными вопросами, которые проверялись в ходе работы, были область применимости ICA анализа ВСР и проблема оптимальной длительности ритмограммы. Для этого проводились расчеты на участках различной длительности от 150 до 1350 ударов сердца (общее время 15 минут). При этом оказалось, что при длине обрабатываемого участка менее 400 точек неверно отображалась низкочастотная компонента разложения за счет

естественной потери низких частот. В то же время, увеличение длины обрабатываемого участка более чем 900 точек, практически не приводило к появлению какой бы, то ни было новой информации, но одновременно резко возрастала коррелированность полученных сигналов. То есть оптимальным будет использование 5 минутного протокола для ICA анализа ВСР.

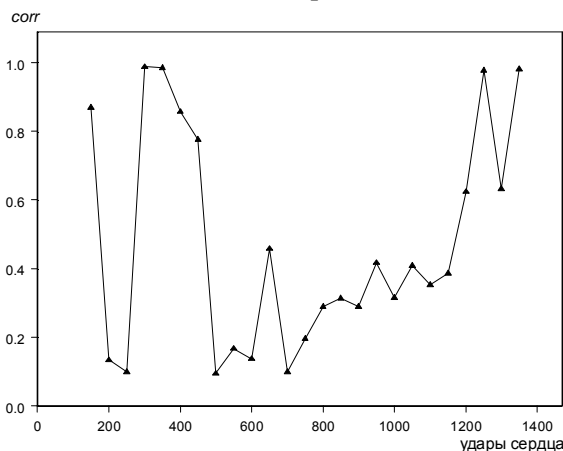


Рис. 5. Корреляции между спектрами для 2-х составляющих

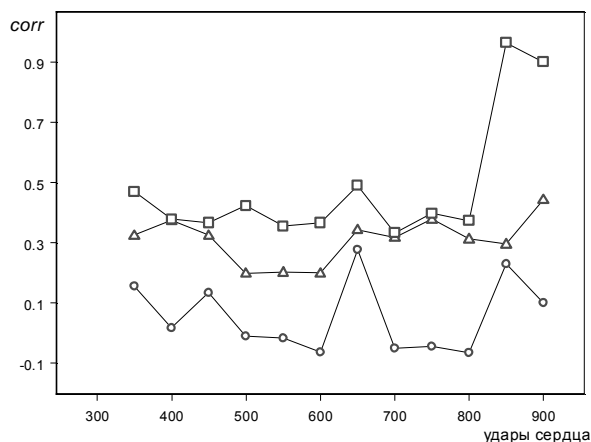


Рис. 6. Корреляции между спектрами для 3-х

ВЫВОДЫ

Формальная математическая процедура ICA анализа ВСР позволяет выделить не бо-

лее трех составляющих в ритмограмме здорового человека, что адекватно отражает физиологические представления об участии систем регуляции в формировании феномена

ВСР. Применение ИСА анализа ВСР оправданно на временных интервалах от 4 до 15 минут. На интервалах до 3 минут оказывается недостаточным количества информации, содержащейся в регистрируемой ритмограмме, для уверенного выделения составляющих. С увеличением времени регистрации более 15 минут происходит значительное осреднение и смешение сигналов от не-

зависимых регуляторных отделов. В обоих случаях выход за указанный временной интервал сопряжен с существенным ухудшением качества выделения формирующих ритмограмму компонент. Оптимальным, с точки зрения качества разделения исходного регистрируемого сигнала ВСР на составляющие, является использование ИСА анализа с 5 минутным протоколом регистрации ВСР.

ЛІТЕРАТУРА

1. Н.И.Яблучанский, А.В. Мартыненко, А.С. Исаева. Основы практического применения неинвазивной технологии исследования регуляторных систем человека. -Харьков «Основа». 2000. 111 с.
2. Aapo Hyvärinen and Erkki Oja, Helsinki University of Technology, Laboratory of Computer and Information Science, Independent Component Analysis, P.O. Box 5400, FIN-02015 Espoo, Finland, 1999.
3. Jutten C. and Herault J.. Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. *Signal Processing*. 1991. Vol. 24. P. 1-10,
4. Comon P. Independent component analysis-a new concept? *SignalProcessing*. 1994. Vol. 36. P. 287-314,
5. Takens F // Lecture Notes in Mathematics. 1981, Vol. 898. Springer-Verlag. P.366-381.
6. Баевский Р.М., Иванов Г.Г.. Вариабельность сердечного ритма: теоретические аспекты и возможности клинического применения/ 2000. 32 с.

НОВА ТЕХНІКА НЕЗАЛЕЖНОГО ФАКТОРНОГО АНАЛІЗУ ВАРІАБЕЛЬНОСТІ СЕРЦЕВОГО РИТМУ

О.В. Мартыненко, Г.С. Антонова, О.М. Єгоренков

Харківський національний університет ім. В.Н. Каразіна

РЕЗЮМЕ

Стаття присвячена розвитку статистичного підходу, заснованого на ідеї розділення реєстрованої композитної ритмограми на незалежні складові, що формуються відповідним регуляторним відділом організму. Формальна математична процедура незалежного факторного аналізу (ІСА) варіабельності серцевого ритму (ВСР) дозволяє виділити не більше трьох складових в ритмограммі здорової людини, що адекватно відображає фізіологічні уявлення про участь систем регуляції у формуванні феномена ВСР. Застосування ІСА аналізу ВСР виправдано на тимчасових інтервалах від 4 до 15 хвилин. Вихід за вказаний часовий інтервал зв'язаний з істотним погіршенням якості виділення формуючих ритмограму компонент. Оптимальним, з погляду якості розділення початкового реєстрованого сигналу ВСР на складові, є використання ІСА аналізу з 5 хвилинним протоколом реєстрації ВСР.

КЛЮЧОВІ СЛОВА: ритмограма, варіабельність серцевого ритму, незалежний факторний аналіз

TECHNOLOGICAL INNOVATION INDEPENDENT COMPONENT ANALYSIS OF HEART RATE VARIABILITY

A.V. Martynenko, A.S. Antonova, A.M. Yegorenkov

V.N. Karazin Kharkiv National University

SUMMARY

The article is devoted to statistical approach that based on method of splitting registered composite rhythmogramm to its independent components, each formed by corresponding regulating units of an organism. Formal mathematical procedure of independent component analysis (ICA) of heart rate variability (HRV) allows to obtain no more than three components forming the rhythmogramm of health person. This fact rationally represents physiological hypotheses about regulation systems taking part in forming HRV phenomenon. Applying ICA to HRV approves itself on timing intervals form 4 to 15 minutes. Breaking the limit of this timing interval causes essential worsening of quality of components forming the rhythmogramm. Optimal application of ICA for splitting initial registered HRV signal into components, according to quality, is using ICA with five-minute HRV registering protocol.

KEY WORDS: rhythmogramm, independent component analysis, heart rate variability